DOCUMENT RESUME

ED 054 199 TM 000 763

AUTHOR Kristof, Walter
TITLE On a Statistic Arising in Testing Correlation.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO RB-71-21
PUB DATE May 71
NOTE 14p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Correlation, Mathematical Applications, *Statistical Analysis, *Statistics, Transformations (Mathematics)

ABSTRACT

The main purpose of this paper consists in deriving means for making statistical inferences about the distribution of u under the conditions of sigma' is equal to sigma" and sigma' is not equal to sigma". An application of the results to coefficient alpha is appended as an illustration. (CK)

RB-71-21

RESEARCH BULLETIN

ED0 54199

ON A STATISTIC ARISING IN TESTING CORRELATION

Walter Kristof

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

TM 000 763

Educational Testing Service
Princeton, New Jersey
May 1971

ON A STATISTIC ARISING IN TESTING CORRELATION

Walter Kristof

Summary

This paper is devoted to the study of a certain statistic, u, defined on samples from a bivariate population with variances σ_{11}, σ_{22} and correlation ρ. Let the parameter corresponding to u be υ. Under binormality assumptions the following is demonstrated. (i) If $\sigma_{11} = \sigma_{22}$, then the distribution of u can be obtained rapidly from the F distribution. Statistical inferences about $\rho = \upsilon$ may be based on F. (ii) In the general case, allowing for $\sigma_{11} \neq \sigma_{22}$, a certain quantity involving u, r (sample correlation between the variables) and υ follows a t distribution. Statistical inferences about υ may be based on t. (iii) In the general case a quantity t' may be constructed which involves only the statistic u and only the parameter υ. If treated like a t distributed magnitude, t' admits conservative statistical inferences. (iv) The F distributed quantity mentioned in (i) is equivalent to a certain t distributed quantity as follows from an appropriate transformation of the variable. (v) Three test statistics are given which can be utilized in making statistical inferences about $\rho = \upsilon$ in the case $\sigma_{11} = \sigma_{22}$. A comparison of expected lengths of confidence intervals for ρ obtained from the three test statistics is made. (vi) The use of the formulas derived is illustrated by means of an application to coefficient alpha.

ON A STATISTIC ARISING IN TESTING CORRELATION[1]

1. Introduction

A paper by Mehta and Gurland (1969) motivated the present study. These authors derived the distribution of a certain statistic, u, defined on samples from a bivariate normal population. They utilized this statistic in estimating the difference of the means of two binormally distributed variables when some of the observations corresponding to one of the variables are missing.

Let X_1, X_2 be binormally distributed variables with variance-covariance matrix $\underline{\Sigma} = \|\sigma_{ij}\|$. Let $\underline{S} = \|s_{ij}\|$ be the matrix of sample second moments. The statistic u is defined as

$$(1) \qquad u = \frac{2s_{12}}{s_{11} + s_{22}} .$$

The corresponding parameter is

$$(2) \qquad \upsilon = \frac{2\sigma_{12}}{\sigma_{11} + \sigma_{22}} .$$

Evidently, $0 \leq \upsilon^2 \leq 1$ and $0 \leq u^2 \leq 1$. Let ρ be the correlation between X_1 and X_2. If $\sigma_{11} = \sigma_{22}$, then $\rho = \upsilon$ and $\hat{\rho} = u$ is the maximum-likelihood estimator of ρ. We will assume that $\upsilon^2 \neq 1$, i.e., $X_2 \neq \pm X_1$ + const.

---

[1]This research was supported by the National Institute of Child Health and Human Development, under Research Grant 1 P01 HD01762.

The main purpose of this paper consists in deriving means for making statistical inferences about υ under the conditions $\sigma_{11} = \sigma_{22}$ and $\sigma_{11} \neq \sigma_{22}$. An application of the results to coefficient alpha will be appended by way of illustration. A more detailed description is contained in the summary.

2. The Case $\sigma_{11} = \sigma_{22}$

Set $\sigma_{11} = \sigma_{22} = \sigma^2$ and make a transformation of variables:

$$(3) \qquad \begin{aligned} Y_1 &= X_1 - X_2 \\ Y_2 &= X_1 + X_2 \quad . \end{aligned}$$

The new variables are again binormally distributed with variance-covariance matrix $\underline{\Sigma}^* = \|\sigma^*_{ij}\|$, say. Explicitly,

$$(4) \qquad \begin{aligned} \sigma^*_{11} &= 2\sigma^2(1 - \rho) \\ \sigma^*_{22} &= 2\sigma^2(1 + \rho) \\ \sigma^*_{12} &= 0 \quad . \end{aligned}$$

Let $S^* = \|s^*_{ij}\|$ be the matrix of new sample second moments. Evidently,

$$(5) \qquad \begin{aligned} s^*_{11} &= s_{11} + s_{22} - 2s_{12} = (s_{11} + s_{22})(1 - u) \\ s^*_{22} &= s_{11} + s_{22} + 2s_{12} = (s_{11} + s_{22})(1 + u) \quad . \end{aligned}$$

We have $\sigma^*_{11}/\sigma^*_{22} = (1 - \rho)/(1 + \rho)$ and $s^*_{11}/s^*_{22} = (1 - u)/(1 + u)$. Hence the quantity

$$(6) \qquad F = \frac{1 - \rho}{1 + \rho} \cdot \frac{1 + u}{1 - u} \ .$$

follows an F distribution with $df_1 = df_2 = N - 1$, N indicating sample size. It would be a simple matter to obtain the distribution of u explicitly from the F distribution by making substitution (6) and using the differential

$$(7) \qquad dF = \frac{1 - \rho}{1 + \rho} \cdot \frac{2du}{(1 - u)^2} \ .$$

This derivation is more elementary and speedier than approaches starting from the Wishart distribution of sample second moments s_{ij}.

The F test provides a familiar means of testing the homogeneity of two random independent sample variances under normality conditions. A hypothesis H_0: $\upsilon = \rho = \rho_0$ is equivalent to homogeneity of $s^*_{11}/(1 - \rho_0)$ and $s^*_{22}/(1 + \rho_0)$, the common parameter being $2\sigma^2$. It follows from (5) that quantity (6) can be employed in testing H_0. Of course, confidence intervals for ρ may also be established by using F.

Rao (1952, p. 226) suggests the likelihood ratio criterion, L, for testing homogeneity of variances when any knowledge as to the possible relationship between the two population variances is missing. In the present context, this amounts to the absence of nontrivial knowledge concerning bounds for ρ. One obtains

$$(8) \qquad L^{\frac{1}{N-1}} = \frac{\sqrt{(1 - \rho_0^2)(1 - u^2)}}{1 - \rho_0 u} \ .$$

For large samples, the quantity $-2\ln L$ is distributed as χ^2 with $df = 1$.

3. The General Case

We admit now $\sigma_{11} \neq \sigma_{22}$. The derivations of the previous section do not generalize to the present case. Instead we will adapt a recent development of the sampling theory of reliability estimation by Kristof (1970).

Using υ as defined in (2), we make the following transformation of variables:

$$(9) \qquad \begin{aligned} Y_1 &= (\sqrt{1-\upsilon} + \sqrt{1+\upsilon})X_1 + (\sqrt{1-\upsilon} - \sqrt{1+\upsilon})X_2 \\ Y_2 &= (\sqrt{1-\upsilon} - \sqrt{1+\upsilon})X_1 + (\sqrt{1-\upsilon} + \sqrt{1+\upsilon})X_2 \quad . \end{aligned}$$

The determinant of the transformation matrix is $4\sqrt{1-\upsilon^2}$. This value is different from zero since $\upsilon^2 \neq 1$. The new variables Y_1, Y_2 are again binormally distributed with variance-covariance matrix $\underline{\Sigma}^* = \|\sigma^*_{ij}\|$, say. Explicitly,

$$(10) \qquad \begin{aligned} \sigma^*_{11} &= 2(1+\sqrt{1-\upsilon^2})\sigma_{11} + 2(1-\sqrt{1-\upsilon^2})\sigma_{22} - 4\upsilon\sigma_{12} \\ \sigma^*_{22} &= 2(1-\sqrt{1-\upsilon^2})\sigma_{11} + 2(1+\sqrt{1-\upsilon^2})\sigma_{22} - 4\upsilon\sigma_{12} \\ \sigma^*_{12} &= 2[2\sigma_{12} - \upsilon(\sigma_{11}+\sigma_{22})] = 0 \qquad \text{by (2).} \end{aligned}$$

Hence the correlation between variables Y_1 and Y_2 is zero.

Let $\underline{S}^* = \|s^*_{ij}\|$ be the matrix of new sample second moments. The quantities s^*_{ij} are obtained from the original sample second moments s_{ij} by substitution of s^*_{ij} and s_{ij} for σ^*_{ij} and σ_{ij}, respectively, in equations (10).

Let us consider a hypothesis H_0: $\upsilon = \upsilon_0$ with $\upsilon_0^2 \neq 1$. Validity of H_0 is equivalent to Y_1 and Y_2 being uncorrelated when υ_0 is used in (9).

The t test provides a familiar means of testing a zero correlation between binormally distributed variables. Under H_0 the quantity

$$(11) \qquad t = \frac{s^*_{12}}{\sqrt{s^*_{11}s^*_{22} - s^{*2}_{12}}}\sqrt{N - 2}$$

follows a t distribution with $df = N - 2$. In terms of the original sample second moments,

$$(12) \qquad t = \frac{u - \upsilon_0}{u\sqrt{1 - \upsilon_0^2}} \cdot \frac{r}{\sqrt{1 - r^2}}\sqrt{N - 2}$$

where $r = s_{12}/\sqrt{s_{11}s_{22}}$ is the sample correlation coefficient.

The derivation of expressions (10) does not depend on distributional assumptions. Binormality was used in giving (11) and its equivalent (12). However, the permutation distribution of a sample correlation coefficient virtually coincides with its normal-theory distribution when the population correlation coefficient is zero (Gayen, 1951). One may therefore expect that the proposed t test (12) "will be reasonably powerful for a wide range of alternatives approximating normality" and, in the absence of binormality, remain "in fact very accurate even for small n [sample size]" (Kendall & Stuart, 1961, pp. 473-476).

Of course, the construction of confidence intervals for υ may also be based on (12) when u and r are given.

Application of formula (12) requires knowledge of both u and r . Let us consider the situation when just u is given and inferences concerning υ are to be made.

Suppose we wish to test a hypothesis H_0: $\upsilon = \upsilon_0$ with $\upsilon_0^2 \neq 1$. Again validity of H_0 is equivalent to Y_1 and Y_2 being uncorrelated when υ_0 is used in (9), $\rho^* = 0$. In analogy to the definition of u we introduce under H_0 the new statistic

$$(13) \qquad u^* = \frac{2s_{12}^*}{s_{11}^* + s_{22}^*} = \frac{u - \upsilon_0}{1 - u\upsilon_0} \quad .$$

Now we have reduced the problem to the following. We are to test a hypothesis H_0^*: $\rho^* = 0$ on the basis of the "observed" value u^*. Unfortunately, the distribution of u^* under H_0^* depends on the unknown ratio $\sigma_{11}^*/\sigma_{22}^*$ as may be seen from formulas (18) and (19) in Mehta and Gurland (1969).

Let us determine the set of all possible sample correlation coefficients r^* that are compatible with a given value u^* when $s_{11}^*, s_{22}^* > 0$. A coefficient r^* is compatible with a given u^* precisely when there are numbers $a,b > 0$ and c such that $u^* = 2c/(a + b)$, $r^* = c/\sqrt{ab}$ and $ab - c^2 \geq 0$. It is a rather simple matter to show that (i) $u^* = 0$ is equivalent to $r^* = 0$, (ii) for $u^* \neq 0$ given, all compatible r^* have the sign of u^* and satisfy precisely $|u^*| \leq |r^*| \leq 1$. Hence the absolutely smallest r^* compatible with a given u^* is u^* itself.

It follows that the quantity

$$(14) \qquad t' = \frac{u^*}{\sqrt{1 - u^{*2}}} \sqrt{N - 2}$$

$$= \frac{u - \upsilon_0}{\sqrt{(1 - u^2)(1 - \upsilon_0^2)}} \sqrt{N - 2}$$

may be used for testing H_0: $\upsilon = \upsilon_0$ when it is treated like a t distributed magnitude with df = N - 2 . This procedure will be conservative in the sense that it will be harder to reject H_0 than when both u and $|r| > 0$ are given and (12) is invoked. We will always have $|t'| \leq |t|$. The relation between t and t' is found to be

$$(15) \qquad t = t'\sqrt{1 + \frac{(s_{11} - s_{22})^2}{4(s_{11}s_{22} - s_{12}^2)}} \quad .$$

The radicand depends on both the difference $s_{11} - s_{22}$ and the determinant $|\underline{S}|$.

4. Further Discussion of the Case $\sigma_{11} = \sigma_{22}$

In the psychometric literature the division of a test into two parallel parts ($\sigma_{11} = \sigma_{22}$, correlation ρ between parts) has received particular interest presumably because the reliability ρ_t of the total test is then a simple monotonic function of ρ alone, namely, $\rho_t = 2\rho(1 + \rho)^{-1}$. Statistical inferences about ρ_t may be based upon statistical inferences concerning ρ .

The assumption $\sigma_{11} = \sigma_{22}$ is evidently equivalent to $\upsilon = \rho$ and, as follows from (10), also to $\sigma^*_{11} = \sigma^*_{22}$. The results of sections 2 and 3 provide us now with at least three different formulas that may be used when inferences about ρ in the case $\sigma_{11} = \sigma_{22}$ are sought. These are:

$$(6) \qquad F = \frac{1 - \rho}{1 + \rho} \cdot \frac{1 + u}{1 - u} \quad , \qquad df_1 = df_2 = N - 1$$

$$(12') \qquad t = \frac{u - \rho}{u\sqrt{1 - \rho^2}} \cdot \frac{r}{\sqrt{1 - r^2}} \sqrt{N - 2} \quad , \qquad df = N - 2$$

$$(14') \qquad t' = \frac{u - \rho}{\sqrt{(1 - u^2)(1 - \rho^2)}} \sqrt{N - 2} \quad , \qquad df = N - 2 \quad .$$

A fourth formula follows from a result first obtained by De Lury (1938) and rederived by Mehta and Gurland (1969). These authors showed that, when $\rho = 0$ and $\sigma_{11} = \sigma_{22}$, the distribution of u based on N pairs of observations is the same as that of r corresponding to $N + 1$ pairs of observations. This result may be applied to u^* as defined in (13) when $\rho^* = 0$, $\sigma^*_{11} = \sigma^*_{22}$ are taken into account. We get

$$(16) \qquad t = \frac{u - \rho}{\sqrt{(1 - u^2)(1 - \rho^2)}} \sqrt{N - 1} \quad , \qquad df = N - 1 \quad .$$

Let us compare formulas (6), (12'), (14') and (16) in terms of the expected length of confidence intervals with equal tail probabilities for ρ at a fixed level when $\sigma_{11} = \sigma_{22}$.

First of all, we have the result that (6) and (16) are equivalent. This is a consequence of the fact that $2^{-1}n^{\frac{1}{2}}(F^{\frac{1}{2}} - F^{-\frac{1}{2}})$ is a t distributed quantity with $df = n$ when F follows an F distribution with $df_1 = df_2 = n$. Hence we need not distinguish between (6) and (16).[2]

Secondly, (16) is uniformly better than (14'). When t is obtained from (16), we have not only always $|t'| \leq |t|$ but the number of degrees of freedom also favors (16).

[2] A proof of the mentioned relation between t and F follows at once from a simple transformation of the variable in the density function of t . The author wishes to thank Professor Ingram Olkin, Department of Statistics, Stanford University, for pointing out that another proof was given earlier by Cacoullos (1965). Cacoullos also reported formula (16).

Thirdly, it has already been said in the previous section that (12') is uniformly better than (14').

Finally, we have to compare (12') and (16). In terms of the starred quantities of the previous section we may write (12') in the form $t = r^*(1 - r^{*2})^{-\frac{1}{2}}(N - 2)^{\frac{1}{2}}$, $df = N - 2$, and (16) as $t = u^*(1 - u^{*2})^{-\frac{1}{2}}(N - 1)^{\frac{1}{2}}$, $df = N - 1$, both when $\rho^* = 0$, $\sigma_{11} = \sigma_{22}$. Under these parameter conditions the previously cited result by De Lury and Mehta/Gurland tells us that the expected length of confidence intervals for ρ will be uniformly shorter when (16) rather than (12') is used.

Thus (16) or, equivalently, (6) is best. Next comes (12') and (14') is last. It is seen that, for N becoming large, the discrepancies between these formulas tend to vanish when $\sigma_{11} = \sigma_{22}$.

On the other hand, when $\sigma_{11} \neq \sigma_{22}$ is admitted, we preferably use (12). Resort will be taken to (14) when r is not available. We must remember that (12) and (14) involve υ instead of ρ in distinction to (12') and (14'). It follows from (15) that $\sigma_{11} \neq \sigma_{22}$ blocks (12) from approaching (14) even when N becomes large.

5. An Application

One possible use of u occurs in psychometric theory. Suppose that a psychological test has been divided into two parts with covariance matrix $\underline{\Sigma} = \|\sigma_{ij}\|$ and binormal score distribution. Coefficient α based upon this division is defined as

$$(17) \qquad \alpha = \frac{4\sigma_{12}}{\sigma_{11} + \sigma_{22} + 2\sigma_{12}} = \frac{2\upsilon}{1 + \upsilon} \quad .$$

This coefficient is a lower bound for the reliability ρ_t of the total test. Novick and Lewis (1967) have shown that it coincides with the reliability precisely when the parts of the total test are essentially τ-equivalent. Consequently, it is generally regarded as a useful quantity in psychometrics. Let $S = \|s_{ij}\|$ be an observed covariance matrix. Then the statistic

$$(18) \qquad \hat{\alpha} = \frac{4s_{12}}{s_{11} + s_{22} + 2s_{12}} = \frac{2u}{1 + u}$$

is the maximum-likelihood estimator of α.

We wish to introduce α and $\hat{\alpha}$ as given in (17) and (18) in expressions (12), (14) and (16). This results in

$$(19) \qquad t = \frac{\hat{\alpha} - \alpha}{\hat{\alpha}\sqrt{1 - \alpha}} \cdot \frac{r}{\sqrt{1 - r^2}}\sqrt{N - 2} \quad , \qquad df = N - 2$$

$$(20) \qquad t' = \frac{\hat{\alpha} - \alpha}{2\sqrt{(1 - \hat{\alpha})(1 - \alpha)}}\sqrt{N - 2} \quad , \qquad df = N - 2$$

$$(21) \qquad t = \frac{\hat{\alpha} - \rho_t}{2\sqrt{(1 - \hat{\alpha})(1 - \rho_t)}}\sqrt{N - 1} \quad , \qquad df = N - 1 \quad .$$

These formulas can be used in testing point hypotheses concerning α or ρ_t, respectively, and in determining confidence intervals for α or ρ_t. Formula (19) was derived earlier by Kristof (1970).

Let us consider a numerical example. From data reported by Lord and Novick (1968, p. 156) one determines for two content-equivalent tests $\hat{\alpha} = 0.9684$, $r = 0.9502$ with $N = 10$. This example was used by Kristof (1970). We wish to establish confidence intervals for α or ρ_t of the composite test at the level $p = 1\%$ with equal tail probabilities. The boundaries of the intervals will be designated $\alpha^{(1)}$ and $\alpha^{(2)}$ or $\rho_t^{(1)}$ and $\rho_t^{(2)}$, respectively.

When the parts are not necessarily regarded as parallel, then the use of (19) is indicated. One obtains $\alpha^{(1)} = 0.9950$, $\alpha^{(2)} = 0.7999$. Now suppose that the observed correlation between the parts was not reported. In this situation we will employ (20). This gives $\alpha^{(1)} = 0.9958$, $\alpha^{(2)} = 0.7632$.

Finally, assume the parts to be parallel. The confidence interval will be determined by means of (21). One arrives at $\rho_t^{(1)} = 0.9952$, $\rho_t^{(2)} = 0.7933$. We see that in a given case a confidence interval calculated by means of (21) need not be shorter than if it were obtained from (19) although, as follows from section 4, the opposite relation holds for the expected lengths of such intervals when $\sigma_{11} = \sigma_{22}$.

It will be noted that the familiar Spearman-Brown formula involving an observed r has never emerged regardless of how a split of a test into two parts is made. Its place is taken by u as is seen from (18).

References

Cacoullos, T. A relation between t and F -distributions. Journal of the American Statistical Association, 1965, 60, 528-531.

De Lury, D. B. Note on correlation. Annals of Mathematical Statistics, 1938, 9, 149-151.

Gayen, A. K. The frequency distribution of the product-moment correlation coefficient in random samples of any size drawn from non-normal universes. Biometrika, 1951, 38, 219-247.

Kendall, M. G., & Stuart, A. The advanced theory of statistics, Vol. 2. New York: Hafner, 1961.

Kristof, W. On the sampling theory of reliability estimation. Journal of Mathematical Psychology, 1970, 7, 371-377.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Mehta, J. S., & Gurland, J. Some properties and an application of a statistic arising in testing correlation. Annals of Mathematical Statistics, 1969, 40, 1736-1745.

Novick, M. R., & Lewis, C. Coefficient alpha and the reliability of composite measurements. Psychometrika, 1967, 32, 1-13.

Rao, C. R. Advanced statistical methods in biometric research. New York: Wiley, 1952.